DOCUMENT RESUME

ED 235 765 HE 016 752


AUTHOR Chickering, Arthur W.
TITLE Quality from the Students' Point of View. AASCU Studies 1983-1.
INSTITUTION American Association of State Colleges and Universities, Washington, D.C.
REPORT NO ISBN-0-88044-071-6
PUB DATE Sep 83
NOTE 23p.; Paper presented at the Annual Meeting of the American Association of State Colleges and Universities (22nd, Nashville, TN, November 2, 1982).
AVAILABLE FROM American Association of State Colleges and Universities, One Dupont Circle, Suite 700, Washington, DC 20036 ($3.00).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Academic Standards; *College Students; Comparative Analysis; *Educational Quality; Evaluation Criteria; Evaluation Methods; *Grading; Higher Education; *Norms; *Student Attitudes; Student Characteristics; *Student Evaluation
IDENTIFIERS *Diversity (Student)



ABSTRACT

Students' perceptions concerning quality in education, the norm-reference approach to student evaluation, and the problem of assessing how much learning has occurred are discussed. To one sample of students, quality meant desirable institutional characteristics and outcomes concerning their own learning and development. Institutions are likely to have difficulty in measuring up to expectations for quality since students have different educational expectations, social and educational background, academic preparation, age, learning styles, and personal development. In addition, the norm-referenced approach to evaluation is not appropriate for a diverse student body. Problems in student evaluation include: each institution and teacher establishes private performance norms, the criteria for performance are seldom clearly articulated, and grades awarded depend primarily upon one student's performance relative to another. Responses to the issue of educational quality and to the underlying complexities of evaluation and grading are recommended. Attention is directed to: the need to recognize different types of competence and knowledge, the fact that standards should vary for different individuals and objectives, and the need to articulate evaluation criteria, methods, and evidence. (SW)



**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED235765

# SCUDIES

1983/1

## Quality from the Students' Point of View

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

AASCU

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CE 016752

American Association of
State Colleges and Universities

# QUALITY FROM THE STUDENTS' POINT OF VIEW

Arthur W. Chickering
Director
Center for the Study of Higher Education
Memphis State University

AASCU STUDIES 1983/1
September 1983

American Association of State Colleges and Universities

©1983 American Association of State Colleges and Universities
One Dupont Circle/Suite 700
Washington, DC 20036

*Library of Congress Cataloging in Publication Data*

Chickering, Arthur W., 1927-
Quality from the students' point of view.

(AASCU studies; 1983/1 (Sept. 1983))
1. Education, Higher–United States–Evaluation. 2. College students–Rating of–United States. 3. Grading and marking (Students) I. Title. II. Series: AASCU studies; 1983/1.
LB2329.5.A43 1983/1 [LA227.3] 378'.01 83-16510
ISBN 0-88044-071-6

*The opinions and recommendations in this book are those of the author and are not intended to represent the official policy of the American Association of State Colleges and Universities.*

# FOREWORD

Those of us involved in higher education, directly and indirectly, have frequently expressed deep frustration over our inability to reach consensus on the basic factors affecting the quality of an educational experience. The frustration mounts because we seem to be convinced that it is essential for consensus to be reached. The growing awareness of the nation's and the world's critical dependence on education fosters a sense of urgency in this matter. Additionally, we have an obligation to ensure effective utilization of the resources available to higher education.

The two concerns above combine to force continuing attention by the American Association of State Colleges and Universities on matters relating to quality of education. As a part of that attention a significant portion of AASCU's 22nd annual meeting in November 1982 was devoted to perceptions and definitions of educational quality. Professor Chickering's essay was presented at that meeting and contains a stimulating point of view which emphasizes the need to relate quality to the diversity of students.

Professor Chickering's presentation brought quickly to mind two other papers I have read in recent years. One of them, by Gordon and Shipman,[1] contained the following: "Concern with human diversity is not new to educators. We have a long history of awareness of individual and group differences. Unfortunately, however, an examination of that history reveals that pedagogical concerns for such differences has been far more obvious in our verbalizations than in our practice. . . . A sizeable body of research has developed around concerns for the individualization of instruction and the exploration of the potential of 'attribute-treatment-interactions'. Yet the range of variance in curriculum design and instructional practice is far less rich than is the diversity to be found in the populations of learners available."

The second,[2] prepared by Clifton Wharton during his presidency of Michigan State University, made the following observation: "A truly excellent pluralistic university does not merely accept its diverse constituency, it responds to that constituency by providing appropriate avenues for the fullest possible realization of actual intellectual potential. Year after year, Michigan State University has reaped the rewards of this responsiveness: in the high school graduate with mediocre grades who then becomes a Rhodes Scholar; in the 'average' freshman who, upon graduation, competes against National Merit Scholars and wins a coveted National Science Foundation scholarship; and in the marginally admissible student who gradually blossoms in the soil of an enriched academic environment and successfully

pursues a graduate program in the field of low-energy physics or microbiology. These actual cases demonstrate the validity of a commitment to provide diverse opportunities for self-development and self-fulfillment."

There is remarkable consistency in the cited papers and the essay presented here by Professor Chickering. Collectively all three papers make very evident that not only is quality *not strained by*, but also that high quality is *not achieved except by* fitting the instructional mode and institutional climate to the educationally relevant characteristics of students. The importance of this view cannot be overemphasized at a time when there appears to be an increasing tendency to ascertain quality by means of some single standard. Moreover, the significance of the view is further magnified by a recognition of the world's need for an immense array of diverse talents, as fully developed as possible.

Professor Chickering's essay merits wide dissemination. Those of us who are involved in one aspect or other of higher education need to be reminded of the challenge facing us. More importantly, however, we need to have before us an agenda of response to the challenge. Professor Chickering seeks to help produce and define the agenda by offering four key suggestions. They deserve, I believe, serious, active attention.

Harold Delaney
Executive Vice President
American Association of State
Colleges and Universities

## ABOUT THE AUTHOR

Arthur W. Chickering is Distinguished Professor of Higher Education and director of the Center for the Study of Higher Education at Memphis State University. After receiving a B.A. from Wesleyan University (1950) and a M.A. from the Graduate School of Education, Harvard University (1951), Chickering completed a Ph.D. in school psychology at Teachers College, Columbia University (1958).

Chickering began his career in higher education as psychology teacher and coordinator of evaluation for the experiment in college curriculum organization at Goddard College from 1959-1965, after working as a secondary school teacher and chairman of the Department of Education at Monmouth College. Chickering then directed the "Project on Student Development in Small Colleges," a four-year study of college influences on student development in thirteen diverse private colleges. During 1970-71, he was a visiting scholar at the Office of Research at the American Council on Education. From 1971 to 1977, he played a major role in creating Empire State College as vice-president for academic affairs.

Chickering is the recipient of the Outstanding Service Award from the National Association of Student Personnel Administrators, a Distinguished Service Citation from the Sex Information and Education Council of the United States, and the Distinguished Contribution to Knowledge Award of the American College Personnel Association. He is a member of the board of trustees of the Council for the Advancement of Experiential Learning, the board of scholars of the Higher Education Research Institute, the editorial board of the *Journal of Higher Education*.

Chickering's major publications include *Education and Identity*, which received the 1969 American Council on Education Book Award for its

---

*Photo courtesy of American Association for Higher Education.*

"outstanding contribution to higher education," *Commuting Versus Resident Students* (1974), *Experience and Learning* (1977), and *Developing the College Curriculum* (with W. Bergquist, D. Halliburton, and J. Lindquist, 1977). In addition, he is the author of numerous journal articles concerning relationships among educational practices, college environments, and student development.

## CONTENTS

FOREWORD ..... iii
by Harold Delaney, AASCU

ABOUT THE AUTHOR ..... v

Introduction ..... 1

Norm-Referenced Evaluation and Student Diversity ..... 4

Evaluation, Grading, and Post-College Performance ..... 7

Assessing Learning Not Status ..... 9

Policies and Practices: Suggestions for Institutional and Professional Development ..... 10

Conclusion ..... 11

NOTES ..... 12

# QUALITY FROM THE STUDENTS' POINT OF VIEW*

**Arthur W. Chickering, Director**
Center for the Study of Higher Education and
Distinguished Professor of Higher Education
College of Education
Memphis State University

## Introduction

What does quality mean from the students' point of view? One rather obvious way to find out is to ask some. We did, and their responses fell into two general categories: (a) outcomes concerning their own learning and development, and (b) desirable institutional characteristics.

For outcomes they said things like, "I want to be competent and knowledgeable in my field." "I want to be able to talk and write about my area clearly." "I want to know what my employment options are in my field." "I want a sense of my own history, how I got to be who I am." "I want a new sense of identity." "I want to develop more respect for myself and my ability to think." "I want to develop increased acceptance and respect for diverse persons, values, ideas, opinions." "I want to develop respect for where my family is and recognition by them that I may be in a different place and that's OK." "I want to become more clear about my own goals and those of the institution."

They also tended to describe institutional conditions that they wanted. "I want respect from the faculty." "I want to be challenged." "I want a chance to learn the way I learn best." I want to be able to express an opinion and not be ridiculed or teased." "I want to be encouraged to ask questions and to see teachers privately." "I want to meet others like myself and share ideas." "I want help in learning how to study and how to learn."

*Prepared for the 22nd AASCU Annual Meeting, Nashville, Tennessee, November 2, 1982. A revised version of this document appeared under the title "Grades: One More Tilt at the Windmill" in the April 1983 *Bulletin* of the American Association for Higher Education.

Some students describe institutional characteristics which would make for a quality experience. "Faculty should know about different kinds of students, what they want, how they learn, what they resent, why they leave." "There should be a handbook for students which tells about courses and not just the content, but how they're taught." "They should have the most experienced and best teachers teach the introductory courses." "Some teachers are better at advising, others better at teaching, others better at research, others better at supervising field experiences. Why don't they let them do what they can do best?" "There should be more access to faculty as resource persons, not just as advisors or course teachers." "There should be better opportunities to connect with professionals in the field I'm interested in." "There should be a peer resource system to help you understand how the college functions and share ideas." "Small voluntary optional study groups could be organized." "How about ongoing seminars with students as presenters and panels of invited reactors."

These outcomes and suggested institutional characteristics don't startle us in their originality. They seem fairly obvious and quite manageable. But somehow not many of our institutions measure up very well to these simple student definitions of quality. Why is that? The basic reason we don't measure up is that those apparently simple statements mean quite different things to different persons. Behind each of those statements there is a complex person quite different in significant ways from many fellow students.

Consider these students:

Bob is 17, high school student with a good record who wants to become a lawyer and is interested in studying history, economics, and political science during his undergraduate career.

Sean, 20, has been working with teenage addicts in a drug treatment center. He dropped out of a four-year college when he found the course requirements and classroom work incompatible with his educational needs, although he's not really certain just what he wants to do.

Charles, 23, a Marine Corps veteran, is a construction worker who also teaches a philosophy course in an experimental school near his home. His goal is to establish an experimental school of his own.

Donald, 34, is an auto worker and has been an elected union officer and bargainer at the national level for 13 years. He wants a degree in labor and economics.

Dianne, 37, has over 200 credit hours in higher education from institutions in the United States, Jamaica, the West Indies, and the British Isles. She is a Registered Nurse and a midwife who is developing curricula for Licensed Practical Nurse programs to upgrade hospital personnel. Eventually she wants to direct such a program in a large medical center.

Alan, 43, has taught electronics for 17 years. He brought 70 credits in English, math, social sciences, and physics, and a contract from Prentice-Hall for a textbook which would be "the culmination of 20 years of work and study in the electronics field." He wants to pursue studies that would help him do the research and writing necessary to complete his book.

Harry, 47, is manager of an internal audit department of a large savings bank. He has 22 years of business experience as well as an A.S. degree in business management and wants to complete a B.S. degree in finance interrupted years ago.

Emily, 56, a home economics aide in the County Department of Social Services, has 111 credits from five different colleges. For 35 years she has worked as a dietitian, nutritionist, adult education teacher, and caseworker aid in home management. She wants to improve her capacity to help disadvantaged persons with low incomes avoid relying on public assistance.

Should the same educational expectations and performance standards be used for these eight students? Or for the thousands of others who present equally complex combinations of backgrounds and purposes? Is there any logic in requiring the same areas of competence or knowledge for all? Could performance criteria be specified that would apply to all these persons, and have positive rather than negative consequences for their learning? It seems unlikely.

The degree program and performance expectations established with 17 year-old Bob, who is academically bright, verbally skilled, and almost totally inexperienced in the worlds of work, of citizenship, of community participation, and who wants to become a lawyer, need to be very different from the programs and expectations set up for Donald autoworker, electronics expert Alan, culturally sophisticated and well-traveled Dianne, wise and socially concerned home economics aide Emily. Most institutions of higher education have not learned to cope with those differences very well. And until we do, we will not measure up very well on quality, either as it may be defined by students, accrediting associations, legislators, or the general public. These wide-ranging student differences in educational background, socioeconomic status, academic preparation, age, learning styles, and personal development are a relatively new phenomenon.

Until the late 1960s each college and university in the United States aimed to serve as carefully selected a group of students as its image and its enrollment pressure would allow. The 1950s and early 1960s probably represented the height of academic meritocracy. The post-World War II baby boom, the GI Bill, and the postwar explosion of interest in higher education created many more applicants than spaces. The opportunity to become increasingly selective became widespread. Admission testing coupled with complex prediction equations became standard practice.

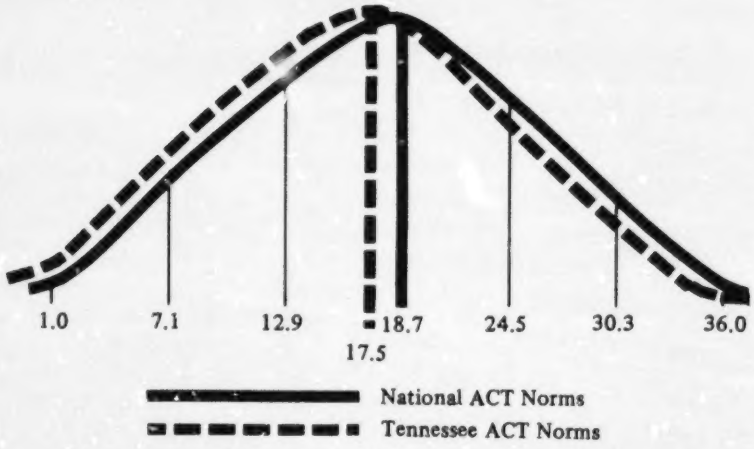
The more these admissions tests and prediction equations evaluated the verbal and academic skills called for by the classroom lecture and written examinations, the more accurate the predictions. High scoring freshmen looked very much like successful graduates, and in many institutions, the performance of some freshmen exceeded the performance of some seniors.

But during the 1960s national and state policies shifted toward a commitment to equal educational opportunity, to providing higher education not only for bright, well-prepared, middle and upper-middle class, mostly white students, but for working class students as well, regardless of race, ethnicity, or pre-existing handicap. This policy shift applied not only to typical college-age students but also to adults of all ages and backgrounds. Higher education is no longer limited to socializing the "elite"; it now aims to educate and train the "masses."

These basic changes from careful selectivity in the service of a meritocratic system to equal opportunity in the service of an egalitarian system have resulted in sharply increased student diversity, especially in publicly supported institutions. They have been accompanied by increased pressures for accountability in providing effective education for these diverse students. Throughout its history higher education has enjoyed the luxury of being able to choose students to fit the college programs; now we must give up that luxury and learn how to design programs to fit the students. Most institutions are still struggling with this new reality. These changed conditions bear most powerfully on our methods and criteria for evaluation. A norm-referenced approach to evaluation has long been standard practice in higher education. But that approach will no longer work in the face of the diverse students to be served.

## Norm-Referenced Evaluation and Student Diversity

Figure 1 illustrates the problem. The top curves give the mean, standard deviations, and distribution curves for the American College Test (ACT) national data and for Tennessee state-supported and two- and four-year institutions. The bottom curves give the same data for five Memphis institutions. Students are clearly sorted by ACT score among these institutions. Each institution establishes its own normative framework for evaluation and grading, and within each, students receive A's, B's, C's, D's, and F's. The result is that students in some institutions receive A's and B's for performances that are graded D or F in another institution or vice versa. Studies of academic variability in colleges across the country carried out in the early sixties showed institutional *means* for graduates which ranged


1.0
7.1
12.9
17.5
18.7
24.5
30.3
36.0
National ACT Norms
Tennessee ACT Norms

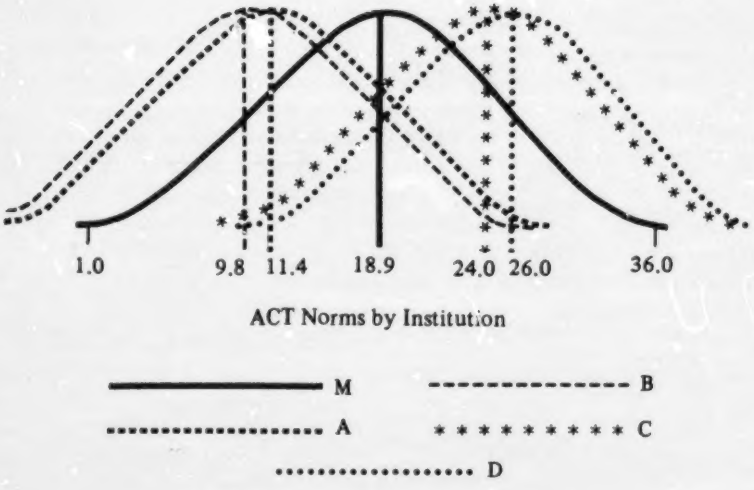

1.0
9.8
11.4
18.9
24.0
26.0
36.0
ACT Norms by Institution
M
B
A
C
D

**Figure 1**

across four standard deviations on measures such as the Graduate Record Examination Area Tests.[3] Given this institutional diversity, the meaning of a given grade becomes parochial in the extreme.

These normative frameworks for evaluation and the grading curves used bear little relationship to any clear or shared indices of performance. Even within each institution those indices are seldom either clearly defined or stable. During the late fifties and early sixties in many institutions, the grading curves remained stable while the ability levels of students increased markedly. Thus in a four-to-six-year period the ACT or Scholastic Aptitude Test (SAT) average for entering students rose by as much as 100 points in those institutions able to be more and more selective, but roughly the same proportions of students received grades A through F. Now that process is being reversed. Entering test scores are dropping at many institutions, but the grading curves remain about the same. In fact, we are worrying about grade inflation while those entering scores are dropping.

In most comprehensive universities distributions of academic ability differ among the different colleges. If differences are significant, does it make sense to assume that the same desired grade distribution should apply to all colleges? And if it is applied, what then do the grades mean in terms of some clear index of performance? But there is often substantial diversity within each college as well. Should the same curves be applied to all majors in all departments? If we adopt a single normative standard across the diverse range of students served by a university, then we will be failing many students who would merit B's and C's at other institutions, and we will be giving A's and B's to students who would merit only C's and D's elsewhere. If we use a normative approach but establish different standards for different college or subgroups within a particular university, we create first, second, and third class degrees for first, second, and third class students. Hardly a direction in which we want to go. These are the boxes we put ourselves in if we rely solely on a norm-referenced approach to evaluation.

K. Patricia Cross, former Educational Testing Service Research Scholar and President of the American Association for Higher Education, and now professor in the Department of Social Policy at Harvard, puts this issue well:

> We know that on any single dimension of human ability half of the students in the nation will be below average. Naive egalitarians have the notion that working with the bottom half will somehow raise them to equal status with the top half. Well, unfortunately, status in the society is

> relative, and on any single measure there will always be a lower half. The only way to reduce the number in the lower half is to expand the number of dimensions along which talent is measured.[4]

Unfortunately, higher education in the past has recognized and tried to foster only a limited array of human abilities. In colleges and universities the term "standards" is seldom used without the modifier "academic." Academic standards as they have been traditionally operationalized have put a premium on reading and writing and on the ability to memorize and manipulate verbal materials. To be sure, underneath this generalization there have been diverse curricular alternatives that recognize the need for additional kinds of competence and knowledge. The sciences and the arts, home economics, physical education and recreation, as they are sometimes taught call for certain mechanical and psychomotor skills; teaching practices in the social sciences sometimes call for certain kinds of interpersonal competence. But basically higher education has been hooked on books, and verbal skills have been the *sine qua non* for effective performance. Now verbal skills are critical for learning and effective performance. So we need to work as hard as we can in that area. But we need to develop criteria for performance in numerous other areas as well and move beyond a norm-referenced approach to evaluating and grading students across our diverse curricula and within our different courses.

## Evaluation, Grading, and Post-College Performance

Our most fundamental assumption is that grades and grade point averages bear some relationship to post-college success, satisfaction, well-being, and the like. Our admission tests and prediction equations have been developed using grade point average as the criterion. The more we selected students like those who achieved high grades the better their chances of success at our institution. Presumably, a better quality graduate was the result. Unfortunately, the evidence—massive evidence from hundreds of studies—does not support those assumptions. The most recent review of the literature that I know of was done by Professor David McClelland, highly respected psychologist in the Department of Psychology and Social Relations, at Harvard University. He said:

> Researchers have in fact had great difficulty demonstrating that grades in school are related to any other behaviors of

importance–other than doing well on aptitude tests. Yet the general public–including many psychologists and most college officials–simply has been unable to believe or accept this fact. It seems so self-evident to educators that those who do well in their classes *must* go on to do better in life that they systematically have disregarded evidence to the contrary that has been accumulating for some time. In the early 1950s a committee of the Social Science Research Council of which I was chairman looked into this matter and concluded that while grade level attained seemed related to future measures of success in life, performance within grade was related only slightly. In other words, being a high school or college graduate gave one a credential that opened up certain higher level jobs, but the poorer students in high school or college did as well in life as the top students .... Such outcomes have been documented carefully by many researchers (cf. Hoyt, 1965) both in Britain (Hudson, 1960) and in the United States. Berg (1970), in a book suggestively entitled *Education and Jobs: The Great Training Robbery*, has summarized studies showing that neither amount of education nor grades in school are related to vocational success as a factory worker, bank teller, or air traffic controller. Even for highly intellectual jobs like scientific researcher, Taylor, Smith and Ghiselin (1963) have shown that superior on-the-job performance is related in no way to better grades in college. The average college grade for the top third in research success was 2.73 (about B-), and for the bottom third, 2.69 (also B-). Such facts have been known for some time.[5]

I don't find these facts surprising, given the problems associated with norm-referenced evaluation discussed above. When each institution and each teacher within each institution establishes parochial and private norms for performance, when the criteria for performance are seldom clearly articulated, when the grades awarded depend primarily upon one student's performance relation to another, when those grades are forced into an arbitrarily agreed-upon curve or distribution, and when student performances are averaged across all the diverse courses and curricula that a program may comprise, then powerful relationships between the resulting GPA and any clearly defined non-college performance area seems unlikely. Back in the fifties Paul Dressel suggested that a grade is an inadequate report of an inaccurate judgment by a biased and variable judge of the extent

to which a student has attained an undefined level of mastery of an unknown proportion of an indefinate amount of material. Unfortunately, I suspect that Paul would think the same point of view was equally accurate today.

## Assessing Learning Not Status

There is one final problem which needs to be recognized. Our discussion thus far has addressed the heavy reliance on norm-references evaluation as the preferred method for evaluating student knowledge and competence at a given point in time: at mid-term, at the end of a semester, in a qualifying or comprehensive exam part way through a graduate program. These examinations do not assess how much *LEARNING* has occurred during the course or graduate program. Most of our evaluation practices assess student status; very few assess the amount of learning which has occurred. We can use simple assessment of students' status concerning a particular area of knowledge or competence if we can safely assume that all students started the course in total ignorance of the subject matter. This assumption, that the student is "tabula rasa," a blank slate to be carefully inscribed by the teacher or institution, may have made sense before the advent of freely accessible print and before the ubiquitous education provided by television and other mass media. And it may have made some sense when we were dealing with a relatively select and homogenous group of 18-year-olds coming straight from high school. But if it made sense in those by-gone days, it no longer makes sense for the increasing numbers of students coming to us of varying ages and backgrounds with many hours before the tube and numerous work and life experiences. Under these conditions relying only on measures of status will tell us little about how much actual learning has occurred. We need more heavy investment in pre-testing and other activities which give us and our students an initial diagnosis of their knowledge and competence at the beginning of a program and course. Without that, our instructional activities will be too difficult for some and too easy for others; we will repeat content already familiar to some and present content so distant from the experiences of others that they cannot profit from it.

We have a substantial challenge simply in finding more effective ways to evaluate and grade the students' status at the end of a course or degree program. If we are to tackle the problem of seriously evaluating the learning which has occurred, we move to a more complex task indeed. The principal justification for the existence of a college or university does not

rest on its capacity simply to provide credentials but on its capacity to create educational environments, teaching practices, and evaluative procedures that result in solid learning for the students to be served. Until we develop the capacity to evaluate that learning and demonstrate our contribution to it, we will be on shaky ground.

## Policies and Practices: Suggestions for Institutional and Professional Development

It is tempting to stop at this point. Like good academics we have indulged our capacity for analysis and criticism. Probably we should leave it there. But there do seem to be some major directions for institutional development and professional development that we might note for whatever they are worth. It seems to us that a sound response to the issue of educational quality and to the underlying complexities of evaluation and grading lie in three major directions.

- Many different types of competence and knowledge must be recognized as legitimate for higher education. Today, there is a widely recognized need to expand the range of student purposes and human potentials served by postsecondary education. Industry, labor, social agencies ask higher education to recognize and contribute to professional/vocational training for employees at all levels. Academics argue for more effective integration of liberal and professional education. To the extent that postsecondary education can respond to these needs society is enriched by diversity, and the possibilities for individual contribution, self-expansion, and satisfaction are increased.

  Publicly supported institutions need to recognize a diverse range of persons as potential college students and a diverse range of purposes as legitimate college pursuits, be they particular intellectual, vocational, or professional skills; knowledge and competence pertinent to various social problems; broader areas of intellectual competence; or more general dimensions of interpersonal competence and personal development. Each college and university needs to help each student clarify his or her major focus for effort and develop high standards for performance in that area. It cannot use a single normative framework as the basis for evaluating performance and learning for the diverse students it serves and the diverse programs it offers.

- Each institution also must recognize that standards should vary from person to person and from purpose to purpose. A general normative approach cannot be used. Our standards will be sound to the degree that the standards established for each combination of individual abilities and purposes are sound. This means that the quality of higher education in our public institutions will depend heavily on their ability to develop teaching practices and educational resources that respond to the individual backgrounds and characteristics of their students. These in turn will depend critically on their ability to develop methods of assessment, evidences of performance, and criteria for evaluation appropriate for the students to be served, for the objectives of the program, and for the gains in knowledge or competence that can reasonably be expected.

- The criteria, the methods, and the acceptable evidence to be used for evaluation must be more clearly articulated. If the range of learning to be recognized is to increase, if variable standards are to be developed appropriate for diverse combinations of students and programs, then we need to become more and more specific about the criteria to be used in certifying acceptable or high-level performance. We need to become much more hard-nosed about the kinds of evidence to be used in making those judgments. Most of us teachers are not used to doing that, and many of us don't know how very well. Most of our evaluations are based on intuitions well honed through practice. Usually our seat-of-the-pants judgments are sound. We do "know a good piece of work when we see one." Or else our evaluations rest on numbers calculated to the third decimal point based on multiple choice exams, leavened by subjective judgments concerning "contributions to class discussion" or the quality of a term paper. But as Dressel observed, most of our grades really result from an inexplicit mixture of (a) student performance relative to some usually inexplicit criteria, (b) the general performance of the class and each student's performance in relation to that general level, (c) the quality of work done by the student in terms of extra papers, readings, lab work, and such, (d) our desire to help student motivation by giving an encouraging grade, and (e) our desire to avoid hassles with students and the administration by giving grades which are not so far out of line with student expectations that tney will make waves. This uncertain and ambiguous mixture, summarized by a single letter or number, will no longer serve. We need to become much more clear about what we are evaluating, how we are evaluating, and what our criteria for judgment seem to be.

- The most important and most difficult challenge is to develop the methods and criteria for evaluating how much learning actually occurs in our various programs and courses. We need to develop diagnostic techniques and measures which can be used as students enter a given program or course so both student and teacher can make sound judgments about the knowledge and competence already available in relation to that which remains to be acquired. This baseline information can guide program planning and instruction. It can provide the basis for making sound judgments about our educational effectiveness, for the quality of our education really rests on the amount of learning we help students achieve, the "value added" to their store of knowledge and competence which occurs as a consequence of their effort and ours.

## Conclusion

When we asked students what quality education meant to them, their first response almost always described significant outcomes concerning learning and personal development. Most of our students do come to us wanting to learn. They also want to be recognized as individuals whose strengths are respected and whose needs for learning are thoughtfully and effectively addressed. Our traditional norm-referenced approaches to evaluation will no longer serve our diverse clientele, nor will they provide either the students or ourselves with sound information concerning how much learning has actually occurred. If we are to respond to increased demands for accountability and if we are to do right by our students, we need to develop procedures for teaching and evaluation that recognize the wide-ranging individual differences in background and purposes brought by our students, and that recognize the diverse types of knowledge and competence required by the complex "learning" and "information" society in which we live.

## NOTES

[1]Gordon, Edmund W. and Stephanie Shipman. "Human Diversity, Pedagogy and Educational Equality." *ERIC/CUE Urban Diversity Series*, Number 59. New York: ERIC Clearinghouse on Urban Education, Teachers College, Columbia University, 1979.

[2]Wharton, Clifton R., Jr. "Excellence in the Pluralistic University." *The Educational Record*, 1973, *54* (Fall), 286-7.

[3]McConnell, T. R. & Heist, P. "The Diverse College Student Population." In N. Sanford, (ed.) *The American College*. New York: Wiley, 1962. pp. 229-236.

[4]Cross, K. P. Personal communication.

[5]McClelland, D. C. "Testing for Competence Rather Than for 'Intelligence'." *American Psychologist*, 1973, 28(1), pp. 1-15.

Additional references cited in the McClelland passage:

Hoyt, D. P. *The Relationship Between College Grades and Adult Achievement, A Review of the Literature*. (Res. rep. No. 7) Iowa City, Iowa: American College Testing Program, 1965.

Hudson, L. "Degree Class and Attainment in Scientific Research." *British Journal of Psychology*, 1960, 51, 67-73.

Berg, I. *Education and Jobs: The Great Training Robbery*. New York: Praeger, 1970.

Taylor, C., Smith, W. R., & Ghiselin, B. "The Creative and Other Contributions of One Sample of Research Scientists," in C. W. Taylor and F. Barron (eds.), *Scientific Creativity: Its Recognition and Development*. New York: Wiley, 1963.

American Association of
State Colleges and Universities

One Dupont Circle/Suite 700
Washington, D.C. 20036
202/293-7070

Allan W. Ostar
President